Remarks

Reconsideration and allowance of the subject patent application are respectfully requested.

The drawings have been changed to, among other things, address the issues identified in the office action. Annotated sheets showing the changes and replacement sheets incorporating the changes are attached in the Appendix to this paper.

As requested, the specification has been amended at page 1, lines 7-9 to provide the application number of the referenced patent application.

Claims 5-8 and 13-16 were objected to because they are alleged to be apparatus claims that depend from method claims. Reconsideration of this objection is respectfully requested. Claims 5-8 and 13-16 describe elements that are configured (or store computer-executable instructions) to perform the method recited in a previous claim. Referencing this previous claim by number rather than by repeating the language of the previous claim is believed to be acceptable. *See, e.g.,* claims 10 and 13 of U.S. Patent No. 6,714,215 and claims 16 and 20 of U.S. Patent No. 6,711,715. Nonetheless, if this objection is maintained, Applicant will re-write the objected to claims to physically incorporate the subject matter of the referenced previous claims. Applicants have amended claims 19 and 27 to address the noted objections.

Claims 1-3, 5-8, 9-11, 13-16, 17-19, 21-24, 25-27 and 29-32 were rejected under the judicially-created doctrine of obviousness-type double patenting based on certain claims of co-pending Application No. 09/658,276. Applicant respectfully requests reconsideration of this rejection in light of the amendments made in this application and the '276 application. If double patenting issues remain, Applicant will take appropriate action such as filing a terminal disclaimer.

Claims 6-8, 14-16, 22-24, and 29-31 were rejected under 35 U.S.C. Section 112, first paragraph, as allegedly failing to provide enablement for an integrated circuit or hardware processing engine. Specifically, the office action states that although the specification is enabling for a mathematical algorithm, it is allegedly not sufficient to enable any person skilled in the art to provide the invention in an integrated circuit or hardware processing engine.

Applicant respectfully traverses this rejection. For example, with the advent of mathematical languages, such as Mathematica®, mathematical equations can be directly converted to computer processing code algorithms. In addition, using Handel-C®, compilable C code can be converted into hardware net-lists. Equations (1)-(4) on page 27, equation (10) on page 29, and equations (11)-(17) on pages 31-32, for example, all use standard algebra and vector representations that are common elements in descriptions either directly translatable into, for example, Mathematica® software, or commonly used in translation to MatLab® software environments. Applicant respectfully submits that a person skilled in the art would be readily able to implement the teachings of this application into hardware engines or integrated circuits and thus withdrawal of the rejection of claims 6-8, 14-16, 22-24 and 29-31 based Section 112, first paragraph, is respectfully requested.

Claims 1, 2, 4-10, 12-18, 20-26 and 28-32 were rejected under 35 U.S.C. Section 102(e) as allegedly being anticipated by Kanevsky *et al.* (U.S. Patent No. 6,421,453). While not acquiescing in this rejection (or in the other rejections discussed below), claims 1, 4, 9, 12, 17, 19, 20, 25, 27 and 28 have been amended. As such, the discussion below is with reference to the amended claims.

Independent claims 1 and 17 each describes, among other things, capturing two or more simultaneous inputs that are responsive to training stimulation; synthesizing the captured inputs; generating a model representation of the synthesized inputs; and using the model to generate outputs in response to real-world stimulation. Independent claims 9 and 25 each describes, among other things, capturing two or more simultaneous inputs that are responsive to training stimulation; synthesizing the captured inputs; generating a model representation of the synthesized inputs; and using the model to generate outputs in response to control command stimulation.

Kanevsky *et al.* describes a method and apparatus that uses a stored sequence of inputs to verify a current sequence of inputs. In particular, Kanevsky *et al.* extracts a sequence of intentional gestures from an individual during a recognition session. This sequence of extracted gestures is compared with a pre-stored sequence of intentional gestures stored during an enrollment session. Recognition/non-recognition is based on the results of the comparison.

Rather than using a sequence of extracted gestures, claims 1, 9, 17 and 25 describe methods and systems in which, among other things, two or more simultaneous inputs (by way of example, simultaneously speaking one's name and signing one's name) are captured and a model is generated that represents a synthesis of these inputs. These features are not disclosed in Kanevsky *et al.* and thus Kanevsky *et al.* cannot anticipate claims 1, 9, 17 and 25.

Generally speaking, prior art incorporating automated decisions using sensor based measurements, such as in a store entrance door opener monitoring a human entering the building, involve a specific set of ***previous measurements*** being incorporated into the human-sensor interaction decision to open the door. Generally, the sensor(s) develop analog voltage signals, which vary with the local environment, such as when a human body-mass gets near a door-opener radar sensor, and reflects back energy that becomes a local measurement of the human presence. This automated decision involves two important concepts: 1) a ***framing time*** is used to collect the measurement data in a digital form and to analyze it to reach a feature-based attribute decision, and 2) a previously determined ***classification threshold*** is used to control the decision process. For each sensory mode an enrollment session is necessary to establish framing times and classification thresholds for integration of a new biometric mode into the decision context of the other modes.

The example systems and methods described in the subject patent application use an entirely new approach to sensor recognition, which requires no predetermined classification processes or frame time measurements for a particular biometric mode. Instead, these example systems and methods establish a continuous interaction of the measured voltage signals in a space, which constructs a model of the human brain thinking and memory processes. Here, the recognition is based on the object-diagramed[1] structures within the constructed space, ***as a memory model of the individual***[2], which was established from a previous observance of the human behavior. With the example systems and methods, a new sensory mode may be

[1] Object diagrams are used as stored representations of the human memory model, and are not references to object-oriented programming.

[2] In Figure 8 of the subject patent application, the response to input stimuli are purposely labeled emotional sensing, not biometric sensing.

introduced without a new enrollment session. The initial enrollment session captures the individual's emotive memory model. This capability is a significant improvement over the Kanevsky *et al.* system, which must depend upon well-known techniques of classification that involves an enrollment for each sensor mode. The example systems and methods use biometric sensing to establish an emotive, thinking and memory model of an individual, establishing a recognition space not disclosed or suggested by Kanevsky *et al.*

In addition, the example systems and methods described in the present patent application can use time frames significantly different from those required for implementing the Kanevsky *et al.* system. The time frames and segmented features appropriate to Kanevsky *et al.* are not adequate to capture the emotive, thinking, and memory model of the subject patent application for the same biometric sensory modes. In other words, the specification of the enrollment and characterization process of Kanevsky *et al.* would not, for example, adequately capture the information necessary to produce the emotive model for verification of an individual described in the subject patent application.

As an example of prior art in biometric recognition, Kanevsky *et al.* involves a set of sensors to monitor a set of intentional gestures, which are compared into a set of previously stored, intentional gesture data, as a set of concurrent inputs, which are processed within a frame time to extract features or attributes. This ***frame time segmentation***, such as with lip motion (Kanevsky *et al.*, col. 6, line 16; Fig. 1), is indexed, and compared as part of the recognition process. The process of merging lip motion into the synchronized speaker recognition measurement data is similar to human lip reading, as that of visually recognizing the lip motion components in speech, concurrent with that of the audio sounds, separated by regions of silence. These specific attributes are derived from measurements, *i.e.*, numbers made every *frame time interval (from time t0 to t1, as 10 msec),* as extracted attributes to align with the prestored, segmented attributes used in recognition (Kanevsky *et al.*, col. 3, lines 6-8; col. 4, lines 27-29). The previously referenced techniques of Kanevsky *et al.* (col. 4, lines 34-38) incorporate sequences of individual attributes, not the sequence of two or more attributes, and hence any improvements in Kanevsky *et al.* are strictly in the alignment of sets of sequenced attributes.

These attributes, taken from two or more sensor sets, are then used in the recognition process (*e.g.*, the lip shape of "O" attribute being aligned with the sound utterance of an open mouth), as a *model of the human vocal system* used in creating the sound (Kanevsky *et al.*, col. 4, line 33). In the same manner, a sequence of events in pen pressure is used as a sequence of attributes extracted during a preset frame time as a *model of the human writing system* (*e.g.*, 5 msec, Kanevsky *et al.*, col. 5, line 15).

The specific recognition comparison is based on a frame-by-frame *feature classification* system (Kanevsky *et al.*, col. 7, line 13-14). Feature classification is a well-established process of comparing feature values, such as lip contour edge numbers (Kanevsky *et al.*, col. 7, line 55), to a known set of previously populated examples of features. A preset threshold (as in a likelihood decision) establishes a probability based decision surface, which is used for acceptance or discard of a segment in the recognition process. The sequence of features in Kanevsky *et al.* uses this in a comparison with a "synchronizer," which aligns the segmented feature sequences (Kanevsky *et al.*, col. 8, line 32), as part of a combined video (lips) and audio (sounds) model (Kanevsky *et al.*, col. 8, line 50) using an "adjustment module" (Kanevsky *et al.*, Figure 3) and a likelihood detector (Kanevsky *et al.*, Figure 3), where the classification threshold is preset in the matching module.

However, the only model used in Kanevsky *et al.* is that of a set of extracted features within a predetermined frame time, being established during a training session. The human vocal attributes are established through the development of the previously determined feature extraction algorithms, and have not used any basis of the simultaneous mental processes of human speech, but rather the framed elements of the vocal track in varying the lip motion (video data) with the changes in the emitted sounds (voice data). The disclosure explicitly lists this as a comparison of synchronized attributes (*see e.g.*, Kanevsky *et al.*, col. 9, line 20; an attribute sequence *aligned to the accuracy of a predetermined frame time*). The degree of this matching is a detailed iteration process, with continuing adjustments using weights as probabilities in a Hidden Markov Modeled process for the alignment of the sequenced attributes (Kanevsky *et al.*, col. 12, line 38 through col. 13, line 8). It is well known that HMM matching, as used in speech

recognition, is based on pre-established probabilities relative to the branching of event sequences, and requires the same frame time in recognition to that of the frame time used in the previously established probabilities. These frame times are typically preset near 5 to 10 msec.

The claims of Kanevsky *et al.* (*see* col. 16, line 53 to col. 17, line 4) are based on this *prestored*, *segmented* by a frame-time, extracted feature set of *attributes*, which are indexed as a *segmented sequence,* and are compared to a measured, segmented sequence of attributes for a match of the sequence index. The Kanevsky *et al.* claims (*see, e.g.,* col. 17, line 46) use the audio and video characteristics of an individual in a joint likelihood function decision of extracted attributes, with step-adjustments in the segment alignment for recognition. This segment adjustment process is referred to as a synchronization process throughout the claims.

The pending claims refer to, among other things, the simultaneous capture of inputs (such as temporally varying inputs) that are synthesized to generate a model representation. Kanevsky *et al.* does <u>not</u> disclose this feature. As noted above, Kanevsky *et al.* uses sequences of individual attributes, not the sequence of two or more attributes. Kanevsky *et al.* <u>synchronizes</u> (or aligns) sets of sequenced attributes, but does not <u>synthesize</u> them. In Kanevsky *et al.*, these synchronized or aligned attributes (*e.g.*, the lip shape of "O" attribute being synchronized or aligned with the sound utterance of an open mouth) are then used in the recognition process. Kanevsky *et al.* does not teach or even remotely suggest synthesizing the inputs to provide, by way of example, a first worldline of linked object diagram exemplars in N-dimensional space based on the inputs and then comparing the worldline to subsequent inputs.

Many of the prior art references identified in the subject patent application utilize sequences of frames and 2-D database recognition structures, with classification trees, which are not that different from the sequenced approach of Kanevsky *et al.* In contrast, the illustrative example embodiments of the subject patent application use concurrent inputs as a simultaneous set of inputs, which have no sequenced segmentation, and, for example, achieve recognition through the ***similarity of the human memory model***, parameterized from a previous, response-stimulation process. The illustrative interface does not use a segmented, framed-time attribute extraction, but rather an automatically measured, atomic element of an object diagram of components in this space. The object diagrams (OD) are the elements of a human memory

model, and the component specificity, becomes a parameterization of the individual human memory at play during the recognition testing.

One aspect the example approach described in this application is in the mathematical complexity of the model representation, comprising an *N-D space, for N>2*, without the need for synchronizing a segmentation of input signals, and in the expanding entropy available for varying the complexity of the model, it is able to match the individuals' complexity of representation in the memory model. The space does not use a likelihood detection scheme, nor an explicit segmentation. Instead, the high dimensional representation space forms a series of projections to lower dimensionality for object component recognition, as being a distinctive human memory process, through the simultaneity of the input interaction across multiple channels. This simultaneous channel interaction is at the full bandwidth of the measurement signals, such as in audio and video are at time intervals of 25 μsec and 0.16 μsec respectively, which are orders of magnitude faster than that of any prior art, due to the absence of a specific frame time. Component recognition is based on N-D structure moments, and the mathematical process of subspace projections simplifies the corrected and verified representation of the space to the observations, as a correctly matching of the inputs to a comparable synthesized output. Much of the mathematical representation is attributed to topographic models in high dimensional spaces, and probabilities are replaced with geometric alignments within these N-D spatial models.

For at least these reasons, Applicant respectfully submits that Kanevsky *et al.* does not anticipate claims 1, 9, 17 and 25. Claims 2, 4-8, 10, 12-16, 18, 20-24, 26 and 28-32 each depends from one of these independent claims and is believed to distinguish over Kanevsky *et al.* because of this dependency and because of the additional patentably distinguishing features contained therein.

Claims 1, 2, 4-10, 12-18, 20-26 and 28-32 were rejected under 35 U.S.C. Section 102(b) as allegedly being "clearly anticipated" by Grossberg (U.S. Patent No. 4,852,018). Grossberg at col. 1, line 44 to col. 2, line 69 discloses a vision system connected to a robotic control, which navigates through the training of associative memory learning for visual location cues, typically implemented in a neural network classifier of feature based information captured within a pre-established frame time. The control of the learning gain is part of the robotic movement

commands. This is a point-to-point movement method for moving a robot along a pathway. Here, Grossberg at col. 2, lines 12-16 describes using simple store and execute position movement commands, similar to many neural- network applications in control theory.

Grossberg does not develop a human memory model, nor use such a model in customization for synthesizing the specific human response as described by way of illustration, not limitation, in the example embodiments of this application. Among other things, there is no synthesis of simultaneous inputs for model construction as specified in the pending claims, and, for example, the representation of real world objects as a worldline in the space of atomic point pathways is never mentioned.

For at least these reasons, Grossberg does not anticipate the subject claims 1, 9, 17 and 25. Claims 2, 4-8, 10, 12-16, 18, 20-24, 26 and 28-32 each depends from one of these independent claims and is believed to distinguish over Grossberg because of this dependency and because of the additional patentably distinguishing features contained therein.

Claims 3, 11, 19 and 27 were rejected under 35 U.S.C. Section 103(a) as allegedly being "obvious" over Kanevsky *et al.* or Grossberg in view of Estes *et al.* (U.S. Patent No. 5,301,284). Estes *et al.* at col. 8, line 53 to col. 11, line 16 describes a means of using an N-D space to represent information as a visual output for human analysis, similar to Starlight® and InSpire® visualization products from PNNL (*e.g.*, the DOE Pacific Northwest National Lab). The image scaling is similar to wavelet scaling used in video compression (*e.g.*, JPEG 2000) and the attribute bit fields are similar to the Machine Vision work of Fujimaka in his retinal neural network application. The OD is just a method of representing the framed features for display. The patent is focused on the visualization of attribute relationships, with layering, which is similar to many graphic-display overlaying techniques, and a typical pseudo-coloring technique of RGB devices. This is more of a database mapping model for related, data element content visualization.

Even assuming for the sake of argument that proper motivation could be identified for combining Estes *et al.* with either Kanevsky *et al.* or Grossberg, Estes *et al.* does not remedy the above-noted deficiencies of the Kanevsky *et al.* and Grossberg patents. Thus, the combination of Estes *et al.* with either of these patents would not result in the subject matter of the rejected

claims. In addition, Estes *et al.* does not disclose or suggest the representation of real world objects as a worldline in, for example, the space of atomic point pathways. Inasmuch as this feature is likewise lacking in Kanevsky *et al.* and Grossberg, the proposed combination is further deficient with respect to claims 3, 11, 19 and 27 in this regard.

New claims 33-52 are added. The subject matter of these new claims is fully supported by the original disclosure and no new matter is added. These claims are believed to be allowable of the documents applied in the office action. For example, claim 35 is directed to a method in which an N-dimensional object space representing a synthesis of simultaneous user inputs is generated and the N-dimensional object space is mapped to one or more M-dimensional sub-spaces to compare the object space representing the synthesis of the simultaneous user inputs to subsequently received simultaneous user inputs. No such method is shown or even suggested in the applied documents and thus claim 35 and claims 36-51 which refer thereto are believed to be allowable. Claim 52 is a system claim corresponding to claim 35 and is likewise believed to be allowable over the applied documents.

The pending claims are believed to be allowable and favorable office action is respectfully requested. Should any issues remain, the Examiner is invited to telephone the undersigned at the number listed below.

Respectfully submitted,

NIXON & VANDERHYE, P.C.

By: Michael J. Shea

Michael J. Shea
Reg. No. 34,725

MJS:mjs
1100 North Glebe Road, Suite 800
Arlington, VA 22201-4714
Telephone: (703) 816-4000
Facsimile: (703) 816-4100